DOCUMENT RESUME

ED 128 665 CG 000 338

AUTHOR Wallace, Phyllis; And Others
TITLE Testing of Minority Group Applicants for Employment. Research Report 1966-7.
INSTITUTION Equal Employment Opportunity Commission, Washington, D.C.
PUB DATE Mar 66
NOTE 35p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Bias; *Culturally Disadvantaged; *Culture Free Tests; *Equal Opportunities (Jobs); Group Intelligence Testing; *Minority Groups; Predictive Validity; *Psychological Tests; Social Discrimination; *Testing

ABSTRACT

The research questions whether many tests used by employers for employee selection do discriminate inadvertantly. Because sub-groups or minority groups tend to be culturally disadvantaged, standardized tests have many shortcomings. When used as selection devices standardized tests may: (1) not provide reliable differentiation in the range of the minority group's scores; (2) use predictive validity different from that for the standardization and validation groups; and (3) use interpretation that is quite dependent upon the cultural background of the majority. To eliminate cultural bias in testing, tests that are culture-free, culture-fair, culture-equivalent, and measure creativity need to be developed. The concepts of using "dual test standards" and "compensatory training" may help create new methods for employee selection. It is concluded that proper testing practices need to be followed by improvement in hiring situations. Recommendations are that: (1) job descriptions be examined and requirements established before tests are used; (2) tests be developed by reputable psychologists; (3) inflexible minimum scores be re-examined; (4) test scores be combined with other data on performance; (5) test scores be validated in the setting used; and (6) re-tests be offered to unsuccessful applicants with culturally disadvantaged backgrounds. (RL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED128665

OFFICE OF RESEARCH AND REPORTS

EQUAL EMPLOYMENT OPPORTUNITY COMMISSION

TESTING OF MINORITY GROUP APPLICANTS FOR EMPLOYMENT

Phyllis Wallace
Beverly Kissinger
Betty Reynolds

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Research Report 1966-7
March, 1966

CG 000 338

ACKNOWLEDGMENTS

This report is the result of intensive research on a highly controversial and complex subject. A number of psychologists have provided us with data from their current studies. We are especially grateful to Commissioner Hernandez who permitted us to examine the testing materials from her files. Dr. Robert Krug, Director of Research for the Peace Corps, made available several of his studies on testing of minority persons. Dr. Philip Ash, Research Assistant to the Vice President of Industrial and Public Relations for Inland Steel; Dr. Joel Campbell of Educational Testing Service; Mr. Howard C. Lockwood of the Lockheed Aircraft Corporation; Dr. Richard Shore, Policy Planning Staff of the Department of Labor; and Dr. Arthur Brayfield, Executive Secretary of the American Psychological Association, have sent us a number of articles. While acknowledging our debt to various scholars, we of course assume full responsiblity for any errors of fact or interpretation.

Phyllis Wallace
Beverly Kissinger
Betty Reynolds

TABLE OF CONTENTS

		Page
Acknowledgments		
I.	Introduction	1
II.	Types of Tests	2
III.	How Tests Discriminate Against Minority Groups	4
IV.	Proposed Solutions to the Problem of Cultural Bias in Testing	9
	A. Variants of Culture-Free and Culture Fair Tests	9
	1. Culture Free Tests	9
	2. Culture Fair Tests	10
	3. Culture Equivalent Tests	10
	B. Creativity Tests	11
	C. Differential Selection Among Applicants from Different Socio-Economic Backgrounds	12
	D. Dual Test Standards and Compensatory Training	14
	E. Intensification of Recruitment	16
	F. Use of Test Scores as Only One Indicator	17
	G. Proper Testing Practices	17
V.	United States As A Model Employer	20
VI.	Recommendations for Testing Guidelines	21
	APPENDIXES	
	Appendix A - The Chronology of the Motorola Case	23
	Appendix B - Glossary of Special Terms	25
	Appendix C - Selected References on Testing	26
	Appendix D - Sources	30

TESTING OF MINORITY GROUP APPLICANTS FOR EMPLOYMENT

I. Introduction

The Motorola case* and the Tower amendment to Title VII of the Civil Rights Act of 1964 Section 703(h)** have dramatized the issue of whether the use of general intelligence tests by employers as selection devices for hiring and promotion deprives Negroes and members of other minority groups of equal employment opportunity. Individuals from culturally disadvantaged*** backgrounds perform less well on these tests on the average than do applicants from middle class environments and consequently may be screened out of training programs and/or excluded from jobs. Differences in culture, in opportunity, and in experience can have a devastating effect on test performance. Since many Negroes, Mexican-Americans, Indians, and lower-class whites have not shared the middle class culture, they may perform in an inferior manner on tests of general intelligence, particularly paper and pencil, but not necessarily on performance for which the tests are supposed to be predictive.

Consistent and significant differences on mean scores are also found between age, sex, educational, and urban-rural groups, but the focus of this report is the effect of testing on the culturally disadvantaged,

*See Appendix A for the chronology of the Motorola case.

** ". . . nor shall it be an unlawful employment practice for an employer to give and to act upon the results of any professionally developed ability test provided that such test, its administration or action upon the results is not designed, intended or used to discriminate because or race, color, religion, sex or national origin." Sec. 703(h)

***See Appendix B.

many of whom are Negroes. This report is not concerned with the willful misuse of tests to discriminate such as giving tests to Negroes but not to whites, or requiring Negroes to achieve higher scores than whites, or failing Negroes regardless of their actual performance. These practices are clearly unlawful. The question to be considered here is whether many "professionally developed ability tests" used by employers to select qualified applicants do in fact discriminate inadvertently.

Authorities in the field of psychological testing have suggested several proposals for mitigating the effects of unintentional types of discrimination against minority groups. We have examined the various proposals and have concluded that careful selection and administering of tests and validation of the testing instrument within an industrial setting, may be the most desirable means to achieve the goal of full utilization of the nation's human resources. The implications of this affirmative conclusion are discussed from the viewpoint of the Equal Employment Opportunity Commission, private employers, and the research psychologists who would have to assume the major responsibility for formulating suitable standards for selection of testing programs.

II. Types of Tests

The major types of tests most commonly used in employee selection are: (1) general intelligence tests, (2) tests of specific intellectual abilities, (3) knowledge and skill tests, (4) measures of dexterity and coordination, and (5) inventories of personality traits.

Intelligence tests such as the Wonderlic, Stanford-Binet, and Otis Quick-Scoring are designed primarily to measure the ability of the individual to understand and to reason with words and numbers. Such tests are most

useful in selection for jobs where learning from and understanding verbal academic material is important.

Specific intellectual abilities tests determine potential for learning certain kinds of work and for solving certain kinds of problems. The tests are not designed to test for a specific job, but to measure the skills for understanding and reasoning with words, numbers and symbols, visualizing of spatial relationships, word fluency, visual speed and accuracy, and creative abilities.

Knowledge and skill tests are usually specific to a job or job family. Knowledge tests are designed to measure the understanding of blueprint reading, electronics, accounting, etc., while skill tests measure one's ability to type, to take dictation, to drive, etc. These tests measure the degree or level of knowledge or skill already attained by candidates at the time of the test.

Dexterity and coordination tests measure speed and accuracy of physical movements. These tests must be very specific to the movements required in the job and are usually constructed by the employer. Examples of such tests are spatial and mechanical abilities, perceptual accuracy, motor abilities.

Personality and interest tests are intended to indicate how a person typically acts and feels, and to determine the type of activities he likes. Tests of this nature have been developed primarily for use in either vocational guidance or clinical use. <u>It is extremely important for a highly trained professional psychologist to evaluate and interpret the results of these tests</u>.

Tests may be further categorized as aptitude versus proficiency. Aptitude tests are designed to measure potential while achievement tests measure skill level at the time of testing.

III. How Tests Discriminate Against Minority Groups

An aptitude test that fails to predict job performance in the same way for both Negroes and whites, or fails to predict job performance at all is not a valid test. If such a test is weighted to differentiate between Negroes and whites, it is similarly invalid and similarly discriminatory. Tests may be held to discriminate in the social sense if they deny equal opportunity for consideration. A test may operate in this manner (a) when scores on it tend to differentiate between identifiable sub-groups, where the sub-grouping itself is not a relevant selection factor, and either (b) scores for the lower group underpredict performance on the job when the standards of the upper-group are applied, or (c) scores on the test do not predict job performance for either group.[1/]

It is known that Negroes on the average do less well on paper and pencil tests than whites. The mean scores for Negroes are lower than the mean scores* for whites on most paper and pencil tests of general ability, intelligence, aptitude, learning ability, or overall ability. The distribution of scores overlap, often considerably, but the average scores differ significantly in most studies.

*Raw scores are converted to norms in order to compare an individual performance with a specific group. See glossary in Appendix B.

More research has been done on the testing of minority group children than workers, but the information which has resulted from this research offers insight into why Negro adults achieve a lower mean score than job applicants from more middle class background. Newton S. Metfessel, Psychologist at the University of Southern California, in his research on children and youth who live in the culture of poverty, found that cultural factors such as home and family structure, personality and social characteristics, learning characteristics, and general school relationships handicap performance on tests.

These children usually come from a home environment where there is such a paucity of objects that the child's conceptual formation development is adversely affected. They also lack curiosity, and this affects both motivational patterns and the development of creative behavior. The culturally disadvantaged child is characterized by weak ego-development, a lack of self-confidence, and a negative self-concept. These conflicting feelings about himself frequently result in exaggerated positive and negative attitudes towards others.[2/]

Many aspects of learning characteristics are affected by the culturally poor background. The culturally disadvantaged typically have a cognitive style which responds more to visual and kinesthetic signals than to oral or written stimuli. Also, these children learn more readily by inductive than deductive approaches. Learning experiences which move from the part to the whole rather than from the whole to the part are

invariably more successful. Significant gaps in knowledge and uneven patterns of learning are typical of this type of background.

Children from the culture of poverty have had little experience in receiving approval for success in learning a task, an assumption on which the school culture is organized. "The cycle of skill mastery which demands that successful experiences generate more motivation to perform which in turn guarantees levels of skill sufficient to prevent discouragement, and so on, may be easily reversed in direction and end the achievement habit prior to its beginning." 3/

In general school relationships and school characteristics, these children from the background of cultural deprivation are placed at a marked disadvantage on standardized tests, which for the most part have been designed to test the white, middle class child. The shortcomings of the standardized tests when they are used with disadvantaged minority groups are discussed below.

A. Reliability of Differentation

Standardized tests may not provide reliable differentation in the range of the minority group's scores. The reliability coefficient for a particular test is strongly affected by the spread of test scores in the group for which the reliability is established. In general, the greater the spread of scores in the reliability samples, the higher the reliability coefficient. For many tests, there is evidence "that children from the lower socio-economic levels tend to have a smaller spread of scores than do children from middle-income families, and such restriction

in the distribution of scores tends to lower reliability so far as differentiation of measurement with such groups is concerned."[4]

B. Predictive Validity

Second, the predictive validity of tests for minority groups may be quite different from that for the standardization and validation groups. Factors which may impair a test's predictive validity are:

1. Test-related factors i.e. test taking skills, anxiety, motivation, speed, understanding of test instructions, degree of item or format novelty, examiner-examinee rapport which may affect test scores but have little relation to the criterion.

2. Complexity of criteria - It is important to recognize the influence of other factors, not measured by tests, which may contribute to criterion success. Since disadvantaged groups tend to do poorly on general intelligence and achievement tests of the paper and pencil type, one should explore background, personality, and motivation of members of such groups for compensatory factors, untapped by the test, which may be related to criterion performance.[5]

While certain aptitude and proficiency tests may have excellent criterion validity for some purposes, even the best of them are unlikely to reflect the true capacity of underprivileged children. They tap abilities that have been molded by the cultural setting. The test content, mode of communication involved in responding to test items, and the motivation needed for making responses are intrinsically dependent upon the cultural context.[6]

C. Validity of Test Interpretation

Third, the validity of the interpretation of tests is strongly dependent upon an adequate understanding of the social and cultural background of the group in question. Sources of error in test interpretation stemming from lack of recognition of the special features of culturally disadvantaged groups are: (1) deviation error - tendency to infer maladjustment from responses which are deviant from the viewpoint of a majority culture, but which may be typical of a minority group. (2) simple determinant error - thinking of the test content as reflecting some absolute or pure trait, process, factor, or construct, irrespective of conditions of measurement or the population being studied. (3) failure barriers - requiring minority group individuals to solve problems with unfamiliar tools.7/

Job applicants from lower socio-economic levels may be characterized in contrast to their middle class counterparts as being less verbal, more fearful of strangers, less confident, less motivated toward scholastic and academic achievement, less conforming to middle class norms of behavior and conduct, less knowledgeable about the world outside their immediate neighborhood. To the extent that these subcultural differences affect test performance adversely, these persons may be denied the opportunity to employment and a more productive contribution to society. Selection instruments often call for responses that are influenced by the culture of the applicant's community or quality of his educational opportunity. Since such tests are "culturally loaded" against persons from a lower socio-economic status, they may

operate as instruments of racial discrimination. The crucial question is whether employers use techniques that unwittingly eliminate persons who might perform satisfactorily on the job. The relationship between test performance and cultural deprivation on the one hand, and job performance on the other, must be investigated for both white and nonwhite job applicants.

IV. Proposed Solutions to the Problem of Cultural Bias in Testing

Most employers defend tests as an efficient device for choosing the most qualified applicants. Where Negro job applicants consistently score significantly below white job applicants a question should be raised about test scores as predictors of job performance. In an employment situation we would like to know whether differences between group means are also associated with performance on the criterion. Do the factors that depress test performance also depress trainability or whatever criterion is to be predicted? Psychologists have suggested ways in which the effect of cultural bias inherent in many aptitude tests can be alleviated for minority group applicants. Few of these proposals have been universally accepted, but most have been discussed in the professional literature on testing of minority groups and the culturally disadvantaged.

A. Variants of "Culture-Free" and "Culture-Fair" Tests

1. Culture-Free Tests

One such proposal is the development of tests which are free of cultural bias in their content and instructions. Dr. Robert Krug, who has written extensively on testing of minority persons, indicates that one of two conditions must be met before a test can be classified as

"culture-free:" either the test items are those which all people of all cultures have had equal opportunity and equal motive to learn, or the test items must possess complete novelty for all people of all cultures.[8] For all practical purposes these two conditions are almost impossible to meet and the idea is often rejected as unfeasible. Howard Lockwood of Lockheed Corporation states that many industrial psychologists agree that even if such a test could be developed, it would be useless in personnel selection. It is impossible, he maintains, to avoid measuring cultural influences, and if they were completely eliminated from all tests, the tests would measure, in essence, nothing.[9]

2. Culture-Fair Tests

Dr. Krug, on the other hand, does not reject the idea entirely. He describes a "culture-fair" test, as a modification of the "culture-free" idea. The assumption underlying the "culture-fair" tests is that there exists a set of test stimuli which are equally appropriate, that is, equal opportunity and motive to learn, for at least two cultural groups.[10] Dr. Paul Schwartz, who headed an AID-sponsored aptitude test development project in West Africa, has done most of the research in this area. A "culture-fair" test or "culture-common test" developed by Schwartz for Nigerian and American children utilized a set of fruits and vegetables which were approximately equal in familiarity to both cultures.

3. Culture-Equivalent Tests

Dr. Schwartz also developed another variant of this concept called "cultural-equivalent" tests, denoting that two tests which are not

identical may, in fact, be equivalent. In this case investigations were undertaken to discover cultural counterparts of tools and machines, cultural manifestations of mechanical principles, and cultural opportunities to acquire information of potential relevance to mechanical training.[11] The argument of cultural equivalence rests on the demonstration that tests constructed in this way have been valid predictors of performance in Westernized training programs in shop mechanics, electrical repair, and the like. Developement of similar tests in this country is impeded by lack of knowledge concerning the culture of southern Negroes, northern slum-dwellers of all races, or any other identifiable sub-groups. Dr. Ash asserts that so-called culture-fair tests do not measure aptitudes or characteristics significantly related to most ordinary measures of job success such as turnover, production or foreman ratings.[12]

B. Creativity Tests

Another approach, adopted by Dr. Newton Metfessel and Professor J. J. Risser, of the University of Southern California involves the use of tests to measure creativity rather than traditional intelligence tests. The latter sample only a relatively small portion of the factors which are involved in intellectual potential and have placed a premium on verbal comprehension and speed of response and emphasize convergent thinking, or the ability to select the one correct answer.[13]

Creativity tests, on the other hand, stress divergent thinking or the ability to create new or original answers. They are, according to Metfessel, more suitable for the testing of the culturally disadvantaged and certain ethnic groups whose command of language is not highly developed.

These tests utilize the most common and familiar of objects in order to sample the testee's ability to recognize problems, and his originality, flexibility, and fluency of thinking. Tasks include suggesting improvements in a familiar device such as a telephone, or thinking of problems that might occur in the use of an object such as a candle. One test requires the subject to list as many uses as he can for a broom handle.[14/]

The tests are scored simply on the number of acceptable answers given by the subject. They seem to be as effective in predicting academic success as traditional intelligence tests and, probably, would be as effective as the latter in predicting job performance.

C. Differential Selection Among Applicants From Different Socio-Economic Ethnic Backgrounds

It has been proposed that, since prediction equations for job performance for most tests currently in use have been based on the performance of whites, different standards (separate test norms, conversion tables, prediction weights, etc.) be employed for Negores and other culturally disadvantaged groups. This approach involves a technique known as the moderator variable. Applicants for a given job

are divided into sub-groups, and selection procedures are applied differentially to members of the two groups. Applicants could be classified, for example, on the basis of a measure of socio-economic status, demographic data (such as percentage of Negroes living in the census tract from which the applicant is applying), and race.

Studies could then be undertaken to determine whether there is, in fact, a differences in the predictive efficiency of job tests as between high and low status groups. Differences in selection procedures for different ethnic groups do not mean a lowering of standards because the standards which count are standards of performance on the job, not the selection standards. Equally qualified persons may be selected from various ethnic groups by applying the standards which are appropriate to each group.[15/]

Lockwood has proposed the use of "cultural exposure" as a moderator variable. Examinees should be grouped homogeneously as to cultural exposure and these groups treated separately in validity studies. Cultural exposure is defined as the material things to which a person has been exposed and the attitudes to which he has been exposed and which he has acquired. Research would lead to a better identification of the culturally disadvantaged and to the utilization of their abilities through a refinement in prediction of training and occupational success.[16/]

A major investigation is under way by Dr. Richard Barrett to determine if the division of applicants into sub-groups improves the accuracy of prediction for members of both groups. If selection is

improved by applying different procedures to the high and low socio-economic groups, then, the more talented would benefit, regardless of race. "It may also happen that dividing the group of applicants on the basis of race may lead to improved accuracy of predictions for members of both races. Such a result has far reaching implications for fair employment practices because failure to treat the two races separately would, if current policies were followed, lead to discrimination against the more talented Negroes." 17/

The overwhelming evidence is that the cultural background of the Negro in America today is so different from that of the white that his performance during the selection process can reasonably be expected to be different. It may be difficult to find an adequate sample of Negroes in most occupations in order to develop separate and suitable prediction equations for them. Lockwood also cautions against the use of a lower minimum score or separate standards of test performance for Negroes since it might tend to perpetuate the idea of race differences or race inferiority.

D. Dual Test Standards and Compensatory Training

The concept of a "dual standard" has some support among psychologists. Ash cites the work of Dr. Kenneth B. Clark of the City University of New York. Clark's work suggests that culturally deprived people who score low on tests may tend to overachieve on the job. In studying the college performance of students who scored low on college entrance tests, Clark found that for students from non-deprived environments, the tests were good predictors, and low college

entrance test scores were accurate indicators of poor grades. On the other hand, students coming from deprived environments did significantly better in college than would have been predicted from the tests.[18]

An experimental training program run by the Federal Department Stores in Detroit, Michigan, indicates that a lowering of required test scores will not necessarily result in a lower quality of job performance. The Federal Department Stores took 16 young people from culturally and economically deprived areas, all of whom had failed standard employment tests and were classified as "unemployable", and put them through a 10-week special training program. All 16 subsequently were employed, 14 at Federal and two elsewhere. The record of performance of all 14 employees at Federal exceeded what was predicted by standard sales aptitude tests. Some exceeded the company's minimum performance standards for new employees by "unbelievable margins."[19]*

Although the Federal Department Stores experiment is considered one of the first of its kind in offering compensatory training for individuals with low test scores, the concept of "double-standard" has had wide acceptance for years in the fairly common practice of maintaining different norms for the sexes. Several popular tests which offer different sex norms are the Bennet, The Wonderlic, the Minnesota Paper Form Board, and the Thurstone Temperament Schedule.

It is generally agreed that some of these sex differences on tests are undoubtedly of environmental origin. Girls, are expected to score lower than boys on tests of mechanical information. It is also

*Re-test results one year later for the ten trainees still employed by Federal showed no significant changes in the scores as a group.

expected that girls will perform less effectively on tasks for which the Mechanical Information test is a predictor. This, however, does not prevent many companies from employing women in manufacturing tasks which require mechanical ability where they perform satisfactorily.[20/]

On the basis of these examples, it appears that a "double-standard" can be justified in some circumstances, though a double standard in job performance and hiring of less qualified applicants is usually rejected as not being effective. If it can be demonstrated that score X for Group A and Score K-k for Group B are associated with identical levels of performance on the job, then an employer might reasonably consider adopting a more flexible attitude toward test scores.[21/]

E. Intensification of Recruitment - While there are significant differences in average performance, there is a considerable overlap in the distribution of test scores of whites and Negroes. It has been proposed, on the basis of this observation, that employers who wish to maintain their present standard of performance on their pre-employment tests, can increase their number of Negro employees by intensifying recruitment among Negroes in order to identify those whose test performance is equal to that of acceptable white applicants. Although this approach has merit in that it could provide employment for Negroes who are qualified but who do not apply for jobs in companies where they assume discrimination is practiced, it is not a solution to the testing problem. It ducks the question of the fairness of tests to those who fail because of

cultural disadvantage, and it will not provide enough additional workers to satisfy present and future labor needs.[22]

F. Use of Test Scores as Only One Indicator - One other practical solution similar in many respects to the "double-standard" is to use test scores as only one indicator among others in the hiring decision, with a clear awareness that, where the applicant has not shared in the predominant middle-class verbal culture, the test score significantly underestimates his potential. A difference of one point more or less cannot be expected to determine if an applicant will fail or succeed on the job. Other personal characteristics such as achievement, motivation, and dependability may be just as significant indicators of successful job performance, and they usually can be identified in each cultural group.

G. Proper Testing Practices - Along with adoption of a more flexible attitude toward test scores, the most immediate improvement can be accomplished by an emphasis by the employer on proper testing practices.

(1) The employer could reconsider the relevance of the qualifications for employment to the specific job tasks required by his company. Many of these requirements are stated in terms of some generalized stereotypes, such as high school graduate, high IQ, or potential to advance to higher level jobs, and are quite extraneous to the requirements of that job. Tests should be professionally chosen to fit the distinctive features of both the industry and the background, education and other characteristics of the successful work force. It is unreasonable to insist

that all lower level workers have potential for supervisory positions. An employer may eventually find that by adopting a more reasonable set of qualifications for each job, he will have access to a considerably larger source of workers who can perform capably and who will present him with fewer problems of employee frustration or labor turnover.

(2) Selection tests should be developed by reputable professional psychologists who are competent in conducting testing programs in an industrial setting.

(3) Pre-employment tests should be administered by personnel who are properly trained not only in the technical details of giving tests, but also in the orientation and handling of people in the testing situation. Members of disadvantaged groups tend to be particularly sensitive to any mannerisms that might be considered antagonistic, sarcastic, or condescending, and test administrators should be aware of this and be able by their behavior to alleviate a certain amount of test anxiety. A personnel manager at a recent testing conference complained that the number of Negro applicants for jobs in his company had fallen off by 80 percent after the company recently instituted a pre-employment testing program.*

(4) A policy of re-testing "failure" candidates may gain for an employer many good employees who otherwise would have been eliminated by the first test. Many candidates, particularly members of minority groups, regard testing as a threatening situation and do not perform as well as they could. A second test would provide a more accurate

*University of Michigan Testing of Minority Group Applicants, January 26, 1966.

indication of the true capability of a person who is less experienced with testing situations and who may have been intimidated by his first experience.

(5) Finally, the most important principle is validation of tests in order to confirm the relationship between test scores and on-the-job performance. There is general agreement that tests should not be used for a group which differs from the validation group. Validity is relative both to the criterion to be predicted and to the group for which the prediction is to be made. Very few employers have validated their testing instruments. In a recent survey by the University of Wisconsin Industrial Relations Research Center, 152 companies which apply testing techniques were canvassed and only 7 percent reported that all their tests had been validated locally against on-the-job performance measures. Nearly 60 percent had validated none of their tests. The remainder reported that some but not all of their tests were validated.[23/]

Dr. Warren Ketcham, University of Michigan psychologist and Vice President of Psychodynamics Research and Associates, has suggested that within company norms should be used exclusively. This only requires that an applicant perform on tests as well as or better than persons who have done or are presently doing the job satisfactorily. The norm tables should then be used to rank applicants as sub-standard, low-average, average, high-average, or superior.[24/]

From recent discussions with research psychologists attached to large industrial concerns, it appears that many companies are developing ability

tests which will measure the essentials required for training or employment, while keeping at a minimum the relevant aspects of culture. For a number of reasons, these findings may never be released for general consumption. One of the responsibilities of the Commission will be to encourage this type of research by the psychological profession. If the purpose of tests is to uncover talent and potential, irrespective of label, surely the Commission could not advocate a more commendable policy.

V. United States As A Model Employer

If the Equal Employment Opportunity Commission establishes basic guidelines on testing of minority group applicants, including a provision on validation of tests, it will require private employers to satisfy certain standards which the United States government, as a civilian employer, for the most part does not meet.*

The U. S. government has set a fine example in its standardized testing program for the military where these tests have been completely validated. Testing in the Armed Forces serves a number of major programs, two of which are (1) to identify the number of personnel required in each skill and professional category, and (2) to identify each individual for training, upgrading, and utilization to his highest potential.

In order to maintain validity, test development activities are mainly serviced by professional job analysts, subject matter specialists, and test psychologists and validated in the working area. This systematic approach is essential to assure that the tests sample specific job functions in

*Of some interest is the fact that the United States Employment Service has recently undertaken a program to develop aptitude measures that can be used to evaluate potential for literacy training, vocational training and occupational potential of the educationally deficient. Much of the research is designed to improve the General Aptitude Test Battery (GATB). 25/

direct proportion to the importance of those functions to the job. As a result, job analysis provides not only a basis for test construction, selection and training, but also a means for increasing productivity and facilitating work.

VI. Recommendations for Testing Guidelines

The following recommendations are designed as a guide to help employers establish objective standards for selection, screening, and promotion of workers. These procedures should ensure that all qualified applicants are given equal opportunity for employment.

1. Job descriptions should be examined and their critical requirements established before tests are selected for screening applicants.

2. Tests used should be those developed by reputable psychologists. Such tests should be administered by professionally qualified personnel who have had training in occupational testing in an industrial setting.

3. Rigidly inflexible minimum scores should be re-examined in light of the considerable research under way on differential selection.

4. Test scores must be considered as only one source of information, and must be combined with other available data on performance such as motivation, leadership and organizational experience, self-sufficiency, and dependability.

5. Tests should be validated within the setting where they will be used. Validation should be for as many separate groups as possible in preference to one large heterogeneous group.

6. It may be advisable for employers who deal with applicants from culturally disadvantaged backgrounds to offer re-tests to candidates who are unsuccessful on their first try, since these people are less familiar with testing situations and may not perform as well as they are able.

APPENDIX A

CHRONOLOGY OF THE MOTOROLA CASE

I. July 15, 1963 - Leon Myart, a Negro, applied for a job as a television phaser and analyzer at the Franklin Park plant of Motorola, Inc. Myart took a five minute intelligence test (General Ability Test No. 10), was interviewed, and was sent home without being told whether he qualified for employment.

II. July 29, 1963 - Failing to receive a job offer, Myart filed a complaint with the Illinois Fair Employment Practices Commission and the President's Committee on Equal Employment Opportunity alleging that his not being hired was due to racial discrimination.

III. January 27-28, 1964 - Hearing of the Motorola case before hearing examiner Robert Bryant of the Illinois Fair Employment Practices Commission.

IV. February 26, 1964 - The hearing examiner directed that Myart be offered a job, that test No. 10 should no longer be used, and that any new test developed in its place should "reflect and equate inequalities and environmental factors among the disadvantaged and culturally deprived groups." He argued that the test had been normed on "advantaged groups" and did not "lend itself to equal employment opportunity to qualify for the hitherto culturally deprived and disadvantaged groups."

V. April 18, May 25, July 14-15, 1964 - Review of the Motorola case before the full Commission.

VI. November 18, 1964 - The Commission issued its unanimous decision, finding that Myart had been denied employment because of his race and while not supporting the order to hire Myart directed that he be compensated one thousand dollars.

VII. April 27, 1965 - Illinois Circuit Court decision on appeal of Motorola. The ruling requiring Motorola to pay Myart one thousand dollars was reversed, but the Commission's findings on discrimination were upheld.

VIII. November 11, 1965 - Case argued before the Illinois Supreme Court.

IX. March 24, 1966 - Illinois Supreme Court reversed the judgment of the circuit court on grounds that the alleged unfair employment practice was not established by a preponderance of the evidence.

APPENDIX B

GLOSSARY OF SPECIAL TERMS

Criterion - A standard that provides a basis for evaluating the validity of a test.

Cultural bias - Propensity of a test to reflect favorable or unfavorable effects of certain types of cultural backgrounds.

Culture-fair test - A test yielding results that are not culturally biased.

Culture-free test - A test yielding results that are not influenced in any way by cultural background factors.

Norms - Statistics that depict the test performance of specific groups. Grade, age, and percentile are the most common types of norms.

Reliability - The degree of consistency, stability, or dependability of measurement afforded by a test.

Validity - The extent to which a test measures the trait for which it is designed, or for which it is being used, rather than some other trait.

Psychological test - An observation of a sample of human behavior made under standard, controlled conditions which results in a linear evaluation called a score.

Culturally disadvantaged - Groups which do not have full participation in American society because of low incomes, substandard housing, poor education, and other "atypical" environmental experiences.

APPENDIX C

Selected References on Testing

1. American Psychological Association, Committee on Scientific and Professional Responsibility, "Social Influences on the Standards of Psychologists," American Psychologist, Vol. 19, 1964, pp.167-173.

2. American Psychologist, Special Issue: Testing and Public Policy. American Psychological Association, Vol. 20, No. 11, November, 1965.

3. Ash, Philip, "Fair Employment Practices Commission Experiences with Psychological Testing," American Psychologist, September 1965, pp.747-798.

4. Ash, Philip, "Race, Employment Tests, and Equal Opportunity," (Presented before Conference of National Association of Inter-Group Relations Officers, Chicago, Illinois, October 21, 1965.)

5. Ash, Philip, "The Implications of the Civil Rights Act of 1964 for Psychological Assessment in Industry," (Presented as part of a symposium, "Legal Issues Which Confront the Psychologist and the Community," 72 Annual APA Convention, Chicago, Illinois September 5, 1965.)

6. Barrett, Richard S., "Differential Selection Among Applicants from Different Socio-Economic Ethnic Backgrounds," Selecting and Training Negroes for Managerial Positions, Princeton, New Jersey, Educational Testing Service, 1965, pp.91-100.

7. Campbell, Joel, "Testing of Culturally Different Groups," Research Bulletin, Princeton, New Jersey, Educational Testing Service, No. RB 64-34, June, 1964.

8. "Can Today's 'Unemployables' Become Tomorrow's Salesmen," (Reprinted with permission from McGraw-Hill, Inc.), New York, New York, American Jewish Committee, March 29, 1965.

9. Chambers, Yolande, "Retraining Program Upsets Test Predictions," Personnel Service, September-October, 1965.

10. Clark, Kenneth B., "Color, Class, Personality, and Juvenile Delinquency," Journal of Negro Education, Vol. 28, 1959, pp.240-251.

11. Coles, Robert, The Desegregation of Southern Schools: A Psychiatric Study, New York, New York, Anti-Defamation League of B'nai B'rith, 1963.

12. Culhane, Margaret M., "Testing the Disadvantaged," The Journal of Social Issues, April, 1964.

13. Dreger, Ralph M., and Miller, Kent S., "Recent Research in Psychological Comparisons of Negroes and Whites in the United States," (Presented at Southeastern Psychological Association, Atlanta, Ga., April 2, 1965.)

14. Dvorak, Beatrice, et al., "New Directions in U. S. Employment Service Aptitude Test Research," Personnel and Guidance Journal, October 1965.

15. Fandell, Todd E., "Testing and Discrimination, Wall Street Journal, April 21, 1964.

16. French, Robert L., "The Motorola Case," The Industrial Psychologist APA Newsletter, Division of Industrial Psychology of the American Psychological Association, Vol. 2, No. 3, August, 1965.

17. Ghiselli, E. E., "A Summary of the Validities of Occupational Aptitude Test," (Presented before the Western Psychological Association, 1965.)

18. Ghiselli, E. E., "Differentiation of Tests in Terms of the Accuracy with which They Predict for a Given Individual," Educational Psychological Measurement, Vol. 20, 1960, pp.675-684.

19. Goslin, D. A., The Search for Ability: Standardized Testing on Social Perspective, New York: Russell Sage Foundation, 1963.

20. "Guidelines for Testing Minority Group Children," (Prepared by a Work Group of the Society for the Psychological Study of Social Issues, Division 9 of the American Psychological Association,) Journal of Social Issues Supplement, Vol. XX, November 2, 1964.

21. Guion, Richard, "Subjectivity in Hiring Standards," Personnel Hiring, McGraw-Hill, 1965, pp.490-493.

22. Katz, I., "Review of Evidence Relating to Effects of Desegregation on the Intellectual Performance of Negroes," American Psychologist, Vol. 19, 1964, pp.381-399.

23. Ketcham, Warren, "Testing Minority Group Applicants," (Prepared for the University of Michigan Bureau of Industrial Relations Personnel Techniques Seminars, January 26, 1966.)

24. Klineberg, Otto, "Negro-White Differences in Intelligence Test Performance: A New Look at an Old Problem," American Psychologist, Vol. 18, 1963, pp.198-203.

25. Krug, Robert E., "Some Suggested Approaches for Test Development and Measurement," (Presented at the symposium, "The Industrial Psychologist, Selection and Equal Employment Opportunity," 1964 Convention of APA, Los Angeles, California, September 4-9, 1964.)

26. Krug, Robert E., "The Problem of Cultural Bias in Selection," Selecting and Training Negroes for Mangerial Positions, Princeton, New Jersey, Educational Testing Service, 1965.

27. Laney, A. R., "Scientific Hiring of Appliance Servicemen," American Gas Association Monthly, January, 1951.

28. Lockwood, Howard C., "Critical Problems in Achieving Equal Employment," (Presented at symposium, "The Industrial Psychologist, Selection and Equal Employment Opportunity," American Psychological Association 1964 Convention, Los Angeles, California, September 1964.)

29. Lockwood,Howard C., "Cultural Exposure and Race as Variables in Predicting Training and Job Success."

30. Lockwood, Howard C.,"Testing Minority Applicants for Employment," (Presented at 1964 Annual Convention of the California State Psychological Association,) Personnel Journal, Vol. 44, July-August 1965, pp.356-360.

31. Lockwood, Howard C., "Progress in Plans for Progress for Negro Managers," (Presented at the Executive Study Conference,New York, New York, November 10, 1964.)

32. Metfessel, Newton S., "Conclusions from Previous Research Findings Which Were Validated by the Research and Evaluation Conducted by the Staff of Project Potential," University of Southern California, 1965

33. Metfessel, Newton S., and Risser, J. J., "Project Potential: Interpretive Guide for the Tests of Creativity," 1965.

34. Ricklefs, Roger, "Jobs and Psychology: Personnel Tests Win Widening Business Use," Wall Street Journal, February, 1965.

35. Rosenberg, Leon A., Rosenberg, Anna M., and Stroud, Michael, "The Johns Hopkins Perceptual Test (The Development of a Rapid Intelligence Test for the Pre-School Child)," April 1966.

36. Runney, George, "Enforcement of Fair Employment Under Civil Rights Act of 1964, "University of Chicago Law Review, Vol. 32, 1965, pp. 430-470.

37. Scioto, Leonard A., and Ryterband, Edward, "Civil Rights and the Industrial Psychologist: A Challenge Not a Threat," The Industrial Psychologist, Vol. 2, 1965, pp. 40-43.

38. Smith, Karl, "Civil Rights and Psychological Testing, "Experimental Cybernetic Foundations of Learning Science, Madison, Wisconsin, University of Wisconsin.

39. Smith, Karl, "Cybernetic Analysis of Personality Assessment Procedures," and "Cybernetic Analysis of Psychological Testing and Test Prediction," Experimental Behavioral Cybernetics, Madison, Wisconsin, University of Wisconsin, June 4, 1965.

40. Smith, Karl, "Proposal for a National Institute of Work Science," Experimental Cybernetic Foundations of Learning Science, Madison, Wisconsin, University of Wisconsin, 1963.

41. Selecting and Training Negroes for Mangerial Positions, Educational Testing Service, Princeton, New Jersey, November, 1965.

42. Spock, Benjamin, "Children and Discrimination," (Reprinted from Redbook), American Jewish Committee, New York, February, 1965.

43. Tumin, Melvin M. (Editor), Race and Intelligence, Anti-Defamation League of B'nai B'rith, New York, 1963.

Motorola Case

44. Circuit Court of Cook County, Illinois
Motorola, Inc. vs. Illinois Fair Employment Practices Commission and Leon Myart (Report of Proceedings)

45. In the Matter of
Leon Myart and Motorola, Inc., State of Illinois, Fair Employment Practices Commission Charge No. 63C-127

46. Supreme Court of Illinois
Motorola, Inc. vs. Illinois FEPC and Leon Myart
(Brief of Plaintiff-Appellant)

47. Supreme Court of Illinois
Motorola, Inc. vs. Illinois FEPC and Leon Myart
(Reply Brief of Plaintiff-Appellant)

48. Supreme Court of Illinois
Motorola, Inc. vs. Illinois Fair Employment Practices Commission and Leon Myart
(Brief and Argument for Illinois Fair Employment Practices Commission, Defendant-Appellee)

49. Supreme Court of Illinois, September Term, A.D. 1965
Motorola, Inc. vs. Illinois Fair Employment Practices Commission and Leon Myart
(Appeal from the Circuit Court)

APPENDIX D

SOURCES*

1. Ash, Philip (5), p.9.
2. Metfessel, Newton (32), p.3.
3. Ibid., p.4.
4. Guidelines (20), p.131.
5. Ibid., p.136.
6. Ibid., p.137.
7. Ibid., p.139-142.
8. Krug, Robert (25), p.6.
9. Lockwood, Howard (30), p.4.
10. Krug (25), p.7.
11. Ibid., p.8.
12. Ash, Philip (4), p.11.
13. Metfessel, Newton (33), p.1.
14. Ibid., p.3.
15. Ash (4), p.13.
16. Lockwood (29), p.4.
17. Selecting and Training Negroes for Managerial Positions (41), p.93.
18. Ash (4), p.12.

*Numbers in parenthesis refer to Appendix C.

19. Merchandising Week (8).

Chambers, Yolande (9)

20. Ash (4), p.5.

21. Krug (25), p.6.

22. Ash (4), p.11.

23. Ibid., (4), p.4.

24. Ketcham (23), p.3.

GSA DC 66-17658